

A regularized NonNegative Canonical Polyadic decomposition algorithm with preprocessing for 3D fluorescence spectroscopy

Jean-Philip Royer*, Nadège Thirion-Moreau*, Pierre Comon[†], Roland Redon[‡], Stéphane Mounier [‡]

*Aix-Marseille Université, CNRS, ENSAM, LSIS, UMR 7296, F-13397 Marseille

Université de Toulon, CNRS, LSIS, UMR 7296, F-83957 La Garde, France, Tel: +33 4 94 14 24 56

{royer, thirion}@univ-tln.fr

[†]GIPSA-Lab, CNRS UMR 5216, Grenoble Campus, BP 46, F-38402 St Martin d'Hères Cédex

pierre.comon@gipsa-lab.grenoble-inp.fr

[‡]PROTEE, Université de Toulon, F-83957, La Garde

{redon, mounier}@univ-tln.fr

Abstract

We consider blind source separation in chemical analysis focussing on the 3D fluorescence spectroscopy framework. We present an alternative method to process the Fluorescence Excitation-Emission Matrices (FEEM): first, a preprocessing is applied to eliminate the Raman and Rayleigh scattering peaks that clutter the FEEM. To improve its robustness versus possible improper settings, we suggest to associate the classical Zepp's method with a morphological image filtering technique. Then, in a second stage, the Canonical Polyadic (CP or Candecomp/Parafac) decomposition of a nonnegative 3-way array has to be computed. In the fluorescence spectroscopy context, the constituent vectors of the loading matrices should be nonnegative (since standing for spectra and concentrations). Thus, we suggest a new NonNegative third order CP decomposition algorithm (NNCP) based on a non linear conjugate gradient optimisation algorithm with regularization terms and periodic restarts. Computer

Nadège Thirion-Moreau is the corresponding author.

This work has been supported by a "PRES euro-mediterranean" grant, and by the European Research Council under the European Programme FP7/20072013, Grant Agreement no.320594, DECODA project.

simulations performed on real experimental data are provided to enlighten the effectiveness and robustness of the whole processing chain and to validate the approach.

Index Terms

Nonnegative tensor factorization (NTF) - nonnegative Canonical Polyadic decomposition (NNCP) - multi-way analysis - preprocessing- 3D fluorescence spectroscopy.

I. INTRODUCTION

Tensor decompositions and more precisely Canonical Polyadic decompositions or Parallel Factor Analysis (Parafac) have been widely used in chemometrics to estimate the concentrations and spectra of chemical species present in solutions. Depending on the considered community, these decompositions have received different names: Canonical Polyadic, Candecomp, CanD, Parafac, but a generally admitted acronym is CP [33], and we shall use it subsequently. This article is focused on the fluorescence spectroscopy framework and deals with the problems of detection, separation and possible classification of several fluorescent compounds (also called fluorophores or fluorochromes) that are dissolved in water samples (or every other type of solvent). These samples can either be prepared in laboratory or collected on-line in different locations (sea, estuaries, rivers or in water and wastewater treatment systems) or acquired at different times. The main purpose of the numerical processing that is then applied is to recover the characteristics of each fluorophore present in the mixtures (*i.e.* the water (or other solvent) samples) for instance their emission and excitation spectra and their concentrations versus time or locations. To reach that aim, Coble has suggested in [15] to use the total luminescence spectra also called “3D fluorescence spectroscopy” or Fluorescence Excitation-Emission Matrices (FEEM). To acquire such signals/images, one uses a spectro-fluorimeter with an excitation source that is continuous versus time. For one given excitation wavelength, an emitted spectrum response can be measured. By sweeping a range of excitation wavelengths, a FEEM can be built. A first really simple but not really efficient way to treat these data was the so-called “peak picking” method which consists of localizing the maximum fluorescence intensities at the excitation and emission wavelengths. Several other methods have been suggested among which are the Orthogonal Projection Approach [21], the Window Factor Analysis [42], the multivariate data analysis method (e.g. Principal Component Analysis) [50], and even blind source separation methods based on Nonnegative Matrix Factorization (NMF) [22]. Yet, the approach which is now largely popularized relies on the CP decomposition of the tensor of FEEM. Such an approach was suggested by Bro [4] and Stedmon [61], in the beginning of the 2000’s. Indeed, a set of several FEEM constitutes a third order tensor which can be modeled by multi-linear decompositions and more precisely by CP decompositions, thanks to a first order approximation of the Beer-Lambert law [60], for weak or low absorbance of the fluorophores. It is the reason

why CP decompositions have been intensively used in the field of environmental data analysis. However, even if an exact fit exists with a known number of terms, the calculation of the CP decomposition actually requires to find the global minimum of a polynomial of degree six or larger (twelve in our case), in a very large number of variables. Such a problem remains numerically very difficult to solve, even if the number of minima is finite. Besides, the model itself might be subject to errors, so that an approximate fit (the best low-rank approximate) is wished to be computed. It is now well known that a best approximate may not always exist [47][17], but it has been proven in [38] that the best low rank nonnegative approximate always exists, and efficient numerical algorithms have been developed for its computation even if proofs of uniqueness are still lacking in the case of a general low-rank approximate. Despite all those problems, one can find many CP algorithms in the literature (see e.g. [4][6][10][13][14][17][20][24][30][34][36][39][46][54][55][56][65][63][68]), exhibiting fluctuating performances versus the level of difficulty (nearly collinear factors when fluorophores possess close spectra, unknown number of fluorophores leading to possible overfactoring problems, etc.).

Yet, before being able to apply CP or NonNegative CP (NNCP) algorithms, the FEEMs have to be preprocessed to eliminate recurring problems that might occur among which are (i) possible non-linear effects (or artifacts such as self-absorption and inner-filter effect when the fluorophores' absorbance become too high to continue to consider a multi-linear model [37]), and (ii) Raman and (iii) Rayleigh scattering peaks that clutter the FEEM and increase the level of difficulty of the problem under consideration. With respect to these latter, subtraction of blanks could be used too. Yet, this procedure might leaves residual scattering peaks that could be problematic when analyzing water samples with low Chromophoric Dissolved Organic Matter (CDOM) concentrations, such as certain coastal and most open ocean waters. This can be the case with other types of solvent too. That is why an alternative technique called Zepp's method was suggested [69].

With the perspective of developing a sensor or a portable monitoring device that could operate *in-situ* and in near real-time to watch environmental data, it remains interesting to develop alternative approaches and to think about new numerical procedures that could be embedded. That is why in this article, our aim is to present a new processing chain for the FEEM. First to eliminate Raman and Rayleigh scatters, we do not use only the classical Zepp's method, which might become too invasive if improper settings are used *i.e.* it might alter the fluorescence signal if the latter is overlapped or passed through by the Raman or Rayleigh scatters. Here, we suggest either to replace it by another preprocessing method based on a morphological image filtering of the FEEM using a well-chosen structuring element [23][58] or to associate these two methods. This procedure requires fewer parameters to be adjusted. It is also helping to make the preprocessing more robust. Then, in a second stage, the CP decomposition of a nonnegative three-way array is computed. Ideally, in fluorescence analysis, the constituent vectors of the three loading matrices (the factors) should be nonnegative since they stand for spectra (excitation and emission) and concentrations. These are intrinsically nonnegative quantities, which justifies the

constraint imposed upon them. There exist different ways to perform constrained minimization [41]. As a result, different CP decomposition algorithms have been developed to take into account this nonnegativity constraint (see e.g. [6][14][46][56][55][9]). This can be ensured in various manners: the simplest one is certainly to consider iterative alternating (minimization) optimization schemes (or Alternating Nonnegative Least Squares (ANLS) approaches) where at each iteration the non-negativity constraint is imposed by a projection on the feasible set. It is the principle of the now well-known NTF-ALS and NTF-HALS algorithms [14]. Otherwise, one could try to solve a constrained optimisation problem (using gradient-based methods like primal-dual interior point methods [43] or direct search methods like Nelder-Mead [44], genetic algorithms and differential evolution [52] or simulated annealing [29]). Another approach consists of approximating the constrained optimization problem by an unconstrained one thanks to a modification of the objective function. In the case of penalty methods (like in [55] for the CP decomposition), a term that prescribes a high cost for violation of the nonnegativity constraints is added, but in the case of barrier methods [45], the term that is added prevents the violation of the constraint since the cost increases to infinity as the unknown parameters approach the boundary of the feasible region. Recently, proximal ANLS approaches [9] as well as active-set-like methods [34] have been considered for NTF too (active set methods were already used in [6]). Finally, another possible approach consists of directly constraining the factors to satisfy the nonnegativity constraint by a change of variables, like a square [56][57] or an exponential [16]: in this case, the tensor decomposition problem is rewritten again as an unconstrained minimization problem.

If such a nonnegativity constraint generally increases computational complexity of the low rank approximation problem, its main advantage is to enable to come back to a well-posed problem [38]. In this article, we opt for one of the NNCP decomposition algorithms that we proposed in [56]. In this approach, the nonnegative nature of the loading matrices is taken into account directly in the problem parameterization using Hadamard products to enforce the searched quantities to be written as squares. Moreover, regularization terms (that might be different from one loading matrix to another one) are also added in order to reinforce certain properties (sparseness or smoothness of the solution). As shown in [67], they improve the robustness of algorithms versus model errors such as a possible overestimation of the tensor rank (which corresponds to the number of fluorescent compounds effectively present in the studied samples, yet, this rank is unknown and it can only be estimated). Without these regularization terms, spurious compounds might appear.

At this stage, to efficiently implement the decomposition, the problem is rewritten as a unconstrained minimization problem of a certain objective function. Thus, the three matrix gradient components (*i.e.* the matrices of the partial derivatives of the objective function versus the unknown loading matrices) have to be derived to be able to use classical optimization algorithms. In this article, we opt for the non linear conjugate gradient (NLCG) algorithm with an Enhanced Line Search (ELS) and periodic restarts. This algorithm offers a good compromise

between efficiency and convergence speed [41][45]. It has the advantage not to require the calculation of Hessian matrices (or their approximations), which might become untractable for big tensors. Moreover, this iterative algorithm is relatively unaffected by the initial solution. Finally, the behavior of the proposed approach is assessed by computer simulations: they are performed on simulated and real experimental data. In this latter case, they correspond to known mixtures of known chemical compounds realized in the PROTEE laboratory to be able to precisely control and validate the results achieved. These results are also compared with those obtained with the method suggested in [3]. We will particularly emphasize the influence of the preprocessing.

This article is organized as follows: some statements about CP decomposition related to fluorescence spectroscopy are recalled in section II. Section III deals with the two preprocessing techniques that are tested and compared. In section IV, we describe the NNCP decomposition algorithm that we suggest to use, and illustrate its good behavior on synthetic data. Section V is dedicated to computer simulations led on experimental data. Comparisons with other existing approaches of the literature are also performed to evaluate pros and cons of each method and better assess their limits. Finally, a conclusion is drawn.

II. PROBLEM STATEMENT: CP DECOMPOSITION OF FLUORESCENT DATA

In this article, we consider two main stages in the processing chain of 3D fluorescence spectroscopy data (e.g. the preprocessing dedicated to the elimination of the most significant types of scattering due to light (Raman and Rayleigh) and the CP decomposition used to separate the contributions of each fluorochrome). The purpose will be to better assess the impact of these two stages on the results achieved.

A. CP decomposition of 3-way arrays

A tensor can be represented by a N -mode array in a chosen basis. We focus on third-order tensors, say $\mathbf{X} \in \mathbb{R}^{I \times J \times K}$. Each entry of \mathbf{X} is denoted by x_{ijk} . Such tensors always admit the following trilinear decomposition, also known as the triadic decomposition [27] of \mathbf{X} into a sum of rank-1 tensors:

$$x_{ijk} = \sum_{f=1}^F a_{if} b_{jf} c_{kf}, \quad (1)$$

where the three involved matrices $\mathbf{A} = (a_{if}) = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_F] \in \mathbb{R}^{I \times F}$, $\mathbf{B} = (b_{jf}) = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_F] \in \mathbb{R}^{J \times F}$, $\mathbf{C} = (c_{kf}) = [\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_F] \in \mathbb{R}^{K \times F}$ are the so-called *loading matrices*, whose F columns are the *loading factors*. F stands for a large enough integer. The minimum F that can be found such that the above equality remains valid is called the tensor rank and the decomposition is then named CP (Canonical Polyadic). The aforementioned model can be rewritten under a more compact form which reads $\mathbf{X} = \llbracket \mathbf{A}, \mathbf{B}, \mathbf{C} \rrbracket$ where

$[\mathbf{A}, \mathbf{B}, \mathbf{C}] = \sum_{f=1}^F \mathbf{a}_f \otimes \mathbf{b}_f \otimes \mathbf{c}_f$ and “ \otimes ” represents an outer product of vectors. It can also be considered under three equivalent matrix form, as follows:

$$\mathbf{X}_{(1)}^{I,KJ} = \mathbf{A}(\mathbf{C} \odot \mathbf{B})^T, \quad \mathbf{X}_{(2)}^{J,KI} = \mathbf{B}(\mathbf{C} \odot \mathbf{A})^T, \quad \mathbf{X}_{(3)}^{K,JI} = \mathbf{C}(\mathbf{B} \odot \mathbf{A})^T, \quad (2)$$

where $(\cdot)^T$ stands for the transpose operator and $\mathbf{X}_{(1)}^{I,KJ}$ is a $I \times KJ$ matrix obtained by unfolding the tensor \mathbf{X} in the first mode (resp. $\mathbf{X}_{(2)}^{J,KI}$ is the $J \times KI$ matrix obtained by unfolding \mathbf{X} in the second and $\mathbf{X}_{(3)}^{K,JI}$ the $K \times JI$ matrix obtained by unfolding in the third mode). Operator \odot stands for the Khatri-Rao product [32]. A standard way to determine the three loading matrices consists of minimizing the following least squares objective function with respect to the three loading matrices:

$$\mathcal{H}(\mathbf{A}, \mathbf{B}, \mathbf{C}) = \frac{1}{2} \|\mathbf{X}_{(1)}^{I,KJ} - \mathbf{A}(\mathbf{C} \odot \mathbf{B})^T\|_F^2 = \frac{1}{2} \|\mathbf{X}_{(2)}^{J,KI} - \mathbf{B}(\mathbf{C} \odot \mathbf{A})^T\|_F^2 = \frac{1}{2} \|\mathbf{X}_{(3)}^{K,JI} - \mathbf{C}(\mathbf{B} \odot \mathbf{A})^T\|_F^2 \quad (3)$$

where $\|\cdot\|_F$ is the Frobenius norm. When performing the CP decomposition, the tensor rank F is assumed to be known and must therefore have been estimated first.

B. Link between 3D Fluorescence Spectroscopy and CP decomposition

When a fluorescent compound dissolved in a solution is excited by a light source at a given frequency, several effects occur: Rayleigh scattering, Raman scattering and fluorescence. At low concentration, the non-linear model predicted by the Beer-Lambert law can be linearized [37]. Thus, the fluorescence intensity rather accurately follows the model below [60][40]:

$$I(\lambda_e, \lambda_f, k) = C \Phi I_o(\lambda_e) \gamma(\lambda_f) \epsilon(\lambda_e) c_k \quad (4)$$

where C is a constant depending on the device, $I_o(\lambda_e)$ is the intensity of the light source, Φ is the fluorescence quantum yield, ϵ denotes the relative absorbance spectrum (sometimes called the excitation spectrum), λ_f is the fluorescence emission wavelength, λ_e stands for the excitation wavelength, γ is the fluorescence relative emission spectrum and c_k is the concentration of the fluorophore in the sample number k . Provided it can be separated from diffusion phenomena, this fluorescence phenomenon allows to determine the contribution to fluorescence and the concentration of a diluted fluorescent chemical component (if ϵ and γ are known), and possibly to identify it, thanks to its fluorescence emission spectrum $\gamma(\lambda_f)$ since database of spectra exist for many fluorescent compounds (see *e.g.* <http://omlc.ogi.edu/spectra/PhotochemCAD/index.html> or <http://www.models.life.ku.dk/domfluor>). It becomes more difficult when the solution contains more than one fluorophore. In such a case, the overall fluorescence intensity is an unknown linear combination of the fluorescence intensities of each of the F fluorophores that are present in the solution. The previous equation becomes:

$$I(\lambda_e, \lambda_f, k) = C I_o(\lambda_e) \sum_{\ell=1}^F \Phi_{\ell} \gamma_{\ell}(\lambda_f) \epsilon_{\ell}(\lambda_e) c_{k,\ell} \quad (5)$$

$c_{k,\ell}$ stands for the concentration of the ℓ -th fluorescent solute in the k -th sample. By sweeping a finite number, say I , of excitation wavelengths, and a finite number, say J , of emission wavelengths and by using K different samples (or mixtures), the fluorescence intensities measurements can be done and stored in a finite third order array of dimensions $I \times J \times K$, denoted by $t_{ijk} = I(\lambda_e(i), \lambda_f(j), k)$.

It becomes obvious, by comparing equations (5) and (1), that when the CP decomposition is unique, one can identify $\epsilon_\ell(\lambda_e(i))$ with a_{il} , $\gamma_\ell(\lambda_f(j))$ with b_{jl} and $\Phi_\ell c_{k,\ell}$ with c_{kl} . Hence, in Eq. (1), the F columns of \mathbf{A} correspond to the excitation spectra of the F fluorophores, the F columns of \mathbf{B} to their fluorescent emission spectra and the F columns of \mathbf{C} to their concentration or contribution and the computation of the CP decomposition yields fluorescent emission spectra of each fluorophore as well as their concentration or contribution and their excitation spectra. The main advantage of such an approach is obviously that there is no need to know in advance what are the components expected to be present in the solution to be able to separate their contribution and then to classify or identify them. Yet, the unknown number of fluorophores, namely F here, remains a difficulty even if different approaches have been suggested to estimate this number (Split Half Analysis [25], COre CONSistency DIAgnostic or CORDONDIA [7], LTMC [28], Threshold-CORCONDIA [18], DIFFIT [64], a convex hull based method [12], SORTe [26], AROFAC2 [35]). In addition, as the quantities that have to be estimated (spectra or concentrations) are inherently non-negative, it has become important to develop CP decomposition algorithms that ensure this non-negativity constraint. Possible applications of such techniques could be the monitoring of the composition of water to be able to early detect the apparition of pollutants in the environment (such as pesticides, Polycyclic Aromatic Hydrocarbons (PAHs) [19], Dissolved Organic Matter (DOM), organic pollutants, soil fulvic acid and even radionuclides [66]) in order to preserve the quality of water [11]. In fact, the fluorescent DOM fractions which are generally observed in surface water, and which provide water quality information, are protein-like - indicator of microbial activity and humic-like - indicator of terrestrial activity. Many studies have shown that fluorescence spectroscopy can characterize natural organic matter [15] and detect different types of aquatic pollutants, like sewage, petroleum products or pesticides [1][8][31][48][49][53]. But another possible application could be the understanding of the interaction between metals salts (like uranyl ion or its complexes) and DOM (or soil fulvic acids) [70] to be able to develop a method or a device to monitor and predict the transport of radioactive metals in aquatic systems.

III. PREPROCESSING: ELIMINATION OF RAMAN AND RAYLEIGH SCATTERING PEAKS

A. Zepp's method

The Zepp's method [69] is a well-known correction method dedicated to the elimination of Rayleigh and Raman scattering peaks. To our opinion, it often leads to good results. It is based on the modeling of the position and spread of the scattering peaks thanks to a set of 4 second degree equations. Then, the scattering peaks

are excised (*i.e.* peak emission ± 10 -15 nm at each excitation wavelength) from the scan data and the excised values are replaced using three-dimensional interpolation of the remaining data by the Delaunay triangulation method. This method thus requires to adjust several parameters (12 for the 4 second degree equations and 8 for the spreading on each side of the 4 curves). However, as illustrated in Fig. 1, in some really difficult cases when the fluorescence peaks and the Raman and Rayleigh scattering peaks are overlapping in FEEM and all the settings of the method are not optimized (here, for example, the spread value is set a little bit too high for one side of one of the 4 curves), this method might become too invasive and the fluorescence peaks might be altered or shifted. To attenuate such a phenomenon, and to make the preprocessing more robust, we suggest to combine this method with a morphological image filtering technique.

B. A technique based on morphological image filtering

Mathematical morphology was introduced by Serra [58] for binary images, at the beginning, using set theory. Its concept was latter extended to gray-scale and color images using both flat and nonflat structuring elements [23][59][62]. In gray-scale mathematical morphology, operations on sets traditionally used in binary mathematical morphology are replaced by operations on domains of functions. The FEEM $I(\lambda_e, \lambda_f, k)$ for all $k = 1 \dots K$ can be viewed as gray-scale images or 3D surfaces whose height is varying over the (λ_e, λ_f) -plane. The principle of mathematical morphology operations is then to use a well-chosen geometric form called the structuring element denoted, here, by $l(\lambda_e, \lambda_f)$ which is used to analyse/modify each of the K studied images/FEEM. In our case, the aim is to remove the Raman and Rayleigh scatters that clutter the FEEM without altering too much the rest of the FEEM. These scatters approximatively take the form of lines of a certain thickness so the mask that plays the role of the structuring element is chosen as a flat rectangle or bar of given sizes (*i.e.* its height is an essential quantity which depends on the thickness of the Raman and Rayleigh scatters (it must be chosen slightly greater than the thickness of the scatters), its width remains small but this quantity plays a less significant role in this case). Another fundamentally important point is that this flat bar must be chosen (approximatively) perpendicular to the Rayleigh and Raman scatters. We opt for a gray-scale “opening” of the FEEM by the aforementioned flat bar because such an operation will decrease the intensity of all bright features provided that their sizes are smaller than those of the structuring element. In fact, the opening of I by l has a geometrical interpretation. It can be seen as “pushing” the structuring element l up from below against the undersurface of I . At each location of the origin of l , the opening is the highest value reached by any part of l as it pushes against the undersurface of I .

A gray scale opening, denoted by \circ , consists of combining two other gray-scale mathematical morphology operations: a gray-scale erosion (\ominus) of the FEEM by the structuring element (bright regions surrounded by dark regions shrink in size with this operation) followed by a gray-scale dilation (\oplus) of the result by the

same structuring element (bright regions surrounded by dark regions grow in size with this operation), *i.e.* $I \circ l = (I \ominus l) \oplus l$. The gray-scale erosion of I by a flat structuring element l at any location (s, t) is defined as the minimum value of the image in the neighborhood of l when the origin of l is at (s, t) . It is mathematically defined as follows:

$$(I \ominus l)(s, t) = \min\{I(s + x, t + y) - l(x, y) | (s + x, t + y) \in \mathbf{D}_I; (x, y) \in \mathbf{D}_l\} \quad (6)$$

For the gray-scale dilatation, the explanation is the same except that the maximum is used instead of the minimum and that the structuring element is reflected about the origin. It is thus defined as:

$$(I \oplus l)(s, t) = \max\{I(s - x, t - y) + l(x, y) | (s - x, t - y) \in \mathbf{D}_I; (x, y) \in \mathbf{D}_l\} \quad (7)$$

\mathbf{D}_I and \mathbf{D}_l respectively standing for the domains of the FEEM I and of the structuring element l .

This operation does not always permit to entirely remove the scattering peaks: residues may be present in the border of the FEEM. That is why we can then either resize the resulting FEEM or apply a second gray-scale “opening” operation by a really small disk of a given radius instead of a bar this second time (this is what is done in all the computer simulations that are presented here).

To illustrate the effect of this preprocessing, we consider the same FEEM as the one studied in the previous subsection. Due to the thickness of Raman and Rayleigh scatters, in this example, we used a flat rectangle-shaped structuring element with a height of 9 pixels (*i.e.* 27 nm) and a width of 1 pixel (*i.e.* 3 nm). After this, we used a disk of radius 3 pixels (*i.e.* 9 nm). The result is displayed in the middle right of Fig. 1. It can be observed that the scatter peaks are not entirely removed but the counterpart is that the fluorescence peaks are less altered than previously as it will be possible to check in section V.

C. Coupling both Zepp and mathematical image filtering to preprocess the FEEM

Denoting by I_{Zepp} the FEEM corrected by the Zepp’s method and I_{Morpho} the FEEM corrected by the mathematical morphology filtering operation described in the previous subsection, the corrected FEEM can be defined as a weighted sum of these two FEEM:

$$I_{\text{corrected}} = \alpha I_{\text{Zepp}} + (1 - \alpha) I_{\text{Morpho}} \quad (8)$$

where α is a real coefficient chosen between 0 and 1. The result of this preprocessing with $\alpha = 0.3$ is displayed in the below of Fig. 1 (this value was tuned empirically while giving a higher weight to mathematical image filtering). It obviously appears as a trade-off between the two previous preprocessing techniques and it enables to be less sensitive to the choice of parameters.

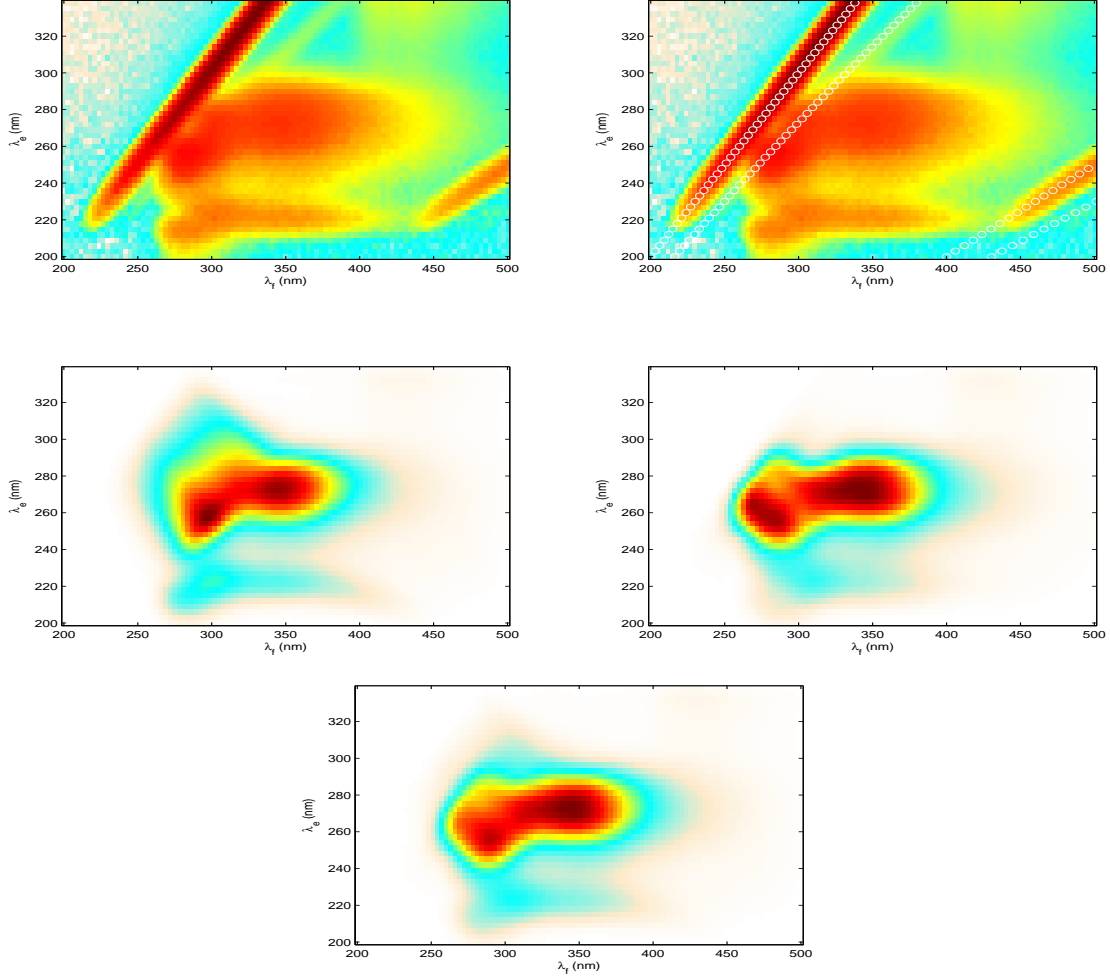


Fig. 1. Above, left: image before processing ; right: modeling of scatter peaks position (\circ) thanks to Zepp's method. Middle, left: image corrected by Zepp's method. Middle, right, image corrected by morphological image processing. Below: image corrected by a combination of both methods (considering $\alpha = 0.3$ in Eq. (8)).

IV. NONNEGATIVE CP DECOMPOSITION OF 3-WAY ARRAYS

A. Suggested approach

Unconstrained CP decomposition leads to an ill-posed problem: the estimation of the loading matrices can become unstable. Such a symptom, known as *degenerate array* [47], occur when collinear columns with opposite signs are estimated in such a way that their contributions cancel each other. This problem has been addressed in various manners taking into account the nonnegativity constraint [6][14][46][34][56]. In [56] we have suggested to constrain the loading matrices to have nonnegative entries by explicitly taking into account their nonnegative

aspect via a parameterization (using a square change of variables) without modifying the cost function. The tensor decomposition problem is thus rewritten as an unconstrained minimization problem. The loading matrices are then defined as $\mathbf{A}' = \mathbf{A} \boxtimes \mathbf{A}$ and the same for \mathbf{B}' and \mathbf{C}' , where \boxtimes is the Hadamard entry-wise multiplication for two matrices with the same dimensions, i.e. $(\mathbf{T} \boxtimes \mathbf{T})_{ij} = t_{ij}t_{ij}$. It leads to a modified cost function:

$$\begin{aligned}\mathcal{F}(\mathbf{A}, \mathbf{B}, \mathbf{C}) &= \mathcal{H}(\mathbf{A} \boxtimes \mathbf{A}, \mathbf{B} \boxtimes \mathbf{B}, \mathbf{C} \boxtimes \mathbf{C}) = \|\mathbf{X}_{(1)}^{I,KJ} - (\mathbf{A} \boxtimes \mathbf{A}) [(\mathbf{C} \boxtimes \mathbf{C}) \odot (\mathbf{B} \boxtimes \mathbf{B})]^T\|_F^2 = \|\boldsymbol{\delta}_{(1)}\|_F^2 \\ &= \|\mathbf{X}_{(2)}^{J,KI} - (\mathbf{B} \boxtimes \mathbf{B}) [(\mathbf{C} \boxtimes \mathbf{C}) \odot (\mathbf{A} \boxtimes \mathbf{A})]^T\|_F^2 = \|\boldsymbol{\delta}_{(2)}\|_F^2 \\ &= \|\mathbf{X}_{(3)}^{K,JI} - (\mathbf{C} \boxtimes \mathbf{C}) [(\mathbf{B} \boxtimes \mathbf{B}) \odot (\mathbf{A} \boxtimes \mathbf{A})]^T\|_F^2 = \|\boldsymbol{\delta}_{(3)}\|_F^2\end{aligned}\quad (9)$$

The calculation of the matrix gradient components [56] leads to:

$$\begin{aligned}\nabla_{\mathbf{A}}\mathcal{F}(\mathbf{A}, \mathbf{B}, \mathbf{C}) &= 4\mathbf{A} \boxtimes ((-\boldsymbol{\delta}_{(1)})[(\mathbf{C} \boxtimes \mathbf{C}) \odot (\mathbf{B} \boxtimes \mathbf{B})]), \\ \nabla_{\mathbf{B}}\mathcal{F}(\mathbf{A}, \mathbf{B}, \mathbf{C}) &= 4\mathbf{B} \boxtimes ((-\boldsymbol{\delta}_{(2)})[(\mathbf{C} \boxtimes \mathbf{C}) \odot (\mathbf{A} \boxtimes \mathbf{A})]), \\ \nabla_{\mathbf{C}}\mathcal{F}(\mathbf{A}, \mathbf{B}, \mathbf{C}) &= 4\mathbf{C} \boxtimes ((-\boldsymbol{\delta}_{(3)})[(\mathbf{B} \boxtimes \mathbf{B}) \odot (\mathbf{A} \boxtimes \mathbf{A})]).\end{aligned}\quad (10)$$

Then, to compute the CP decomposition, an optimization algorithm has to be chosen. In [56], we studied different deterministic “descent” algorithms among which are Quasi-Newton methods (BFGS and DFP algorithms [41][51]) and the nonlinear conjugate gradient (NLCG) algorithm. This latter has been shown to offer the best compromise between efficiency and convergence speed. Moreover, as it does not require the calculation of Hessian matrices (or their approximations) like Quasi-Newton methods, it can still be used in problems involving large tensors. Subsequently, we will use it in the following. Using (10), the two following $(I + J + K)F \times 1$ vectors are built:

$$\mathbf{g}^{(k)} = \begin{pmatrix} \text{vec}\{\nabla_{\mathbf{A}}\mathcal{F}(\mathbf{A}^{(k)}, \mathbf{B}^{(k)}, \mathbf{C}^{(k)})\} \\ \text{vec}\{\nabla_{\mathbf{B}}\mathcal{F}(\mathbf{A}^{(k)}, \mathbf{B}^{(k)}, \mathbf{C}^{(k)})\} \\ \text{vec}\{\nabla_{\mathbf{C}}\mathcal{F}(\mathbf{A}^{(k)}, \mathbf{B}^{(k)}, \mathbf{C}^{(k)})\} \end{pmatrix}, \quad \mathbf{x}^{(k)} = \begin{pmatrix} \text{vec}\{\mathbf{A}^{(k)}\} \\ \text{vec}\{\mathbf{B}^{(k)}\} \\ \text{vec}\{\mathbf{C}^{(k)}\} \end{pmatrix}\quad (11)$$

where the $\text{vec}\{\cdot\}$ operator applied on a given matrix stacks its columns into a column vector.

When applied to non quadratic problems, the non linear conjugate gradient (NLCG) can take several forms [41]. Here, we focus on line search methods. The NLCG algorithm (using restarts) reads:

Step 1. Given $\mathbf{x}^{(0)}$ (Eq. (11)), compute $\mathbf{g}^{(0)}$ (Eq. (11)) and set $\mathbf{d}^{(0)} = -\mathbf{g}^{(0)}$

Step 2. For $k = 0, 1, \dots, n - 1$

$$\begin{cases} \mathbf{x}^{(k+1)} &= \mathbf{x}^{(k)} + \mu^{(k)}\mathbf{d}^{(k)} \\ \mathbf{d}^{(k+1)} &= -\mathbf{g}^{(k+1)} + \beta^{(k)}\mathbf{d}^{(k)} \end{cases}\quad (12)$$

Step 3. If $k = n - 1$ replace $\mathbf{x}^{(0)}$ by $\mathbf{x}^{(n)}$ and go back to Step 1 (*i.e.* restart).

In exact line search methods, two expressions for β are classically used: the Fletcher-Reeves (β_{FR}) and the Polak-Ribière (β_{PR}) formula [51] (we use β_{FR}). Finally, the restarting aspect of the NLCG algorithm is fundamental for the global convergence analysis [41][51] since in general one cannot guarantee that the directions $\mathbf{d}^{(k)}$ are effectively descent directions. But since a pure steepest descent step is taken at each restart by setting $\mathbf{d}^{(k)} = -\mathbf{g}^{(k)}$ the local convergence is assured. It also leads to better performance and the algorithm converges faster. In our case, this “restart” is normally performed every $n - 1 = (I + J + K)F$ iteration but it can also be forced if $|\mathbf{g}^{(k)}|$ becomes too important. At each iteration, matrices \mathbf{B} and \mathbf{C} are rescaled to have unit- l_1 norm columns while this condition is not imposed on matrix \mathbf{A} : this choice is made to allow matrix \mathbf{A} to have columns of zeros in the case of overfactoring. In case of rescaling, the scale factor is then pulled into \mathbf{A} .

B. Addition of regularization terms

The previous objective function \mathcal{F} can be modified by adding regularization terms whose aim is to impose boundedness on the solution and/or to enforce another specific property on the solution such as smoothness, sparsity or uncorrelatedness. The modified objective function finally reads:

$$\mathcal{G}(\cdot) = \mathcal{F}(\cdot) + \alpha_A \|\mathbf{A} \boxminus \mathbf{A}\|_{\bullet} + \alpha_B \|\mathbf{B} \boxminus \mathbf{B}\|_{\bullet} + \alpha_C \|\mathbf{C} \boxminus \mathbf{C}\|_{\bullet} \quad (13)$$

where α_A , α_B and α_C are nonnegative regularization parameters to impose boundedness on the solution. In (13), different kinds of norms $\|\cdot\|_{\bullet}$ can be considered. The standard Tikhonov l_2 -norm regularization ($\|\cdot\|_{\bullet} = \|\cdot\|_F^2$) is meant to enforce smoothness of the solution while the l_1 -norm regularization ($\|\mathbf{C}\|_{\bullet} = \|\mathbf{C}\|_1 = \sum_{i,j} |c_{ij}|$) is meant to enforce sparsity of the solution. The gradient components given in (10) are simply replaced by:

$$\begin{aligned} \nabla_{\mathbf{A}} \mathcal{G}(\cdot) &= \nabla_{\mathbf{A}} \mathcal{F}(\cdot) + \alpha_A \star & \text{where } \star &= 2\mathbf{A} \quad \text{or} \quad 4\mathbf{A} \boxminus \mathbf{A} \boxminus \mathbf{A}, \\ \nabla_{\mathbf{B}} \mathcal{G}(\cdot) &= \nabla_{\mathbf{B}} \mathcal{F}(\cdot) + \alpha_B \star\star & \text{where } \star\star &= 2\mathbf{B} \quad \text{or} \quad 4\mathbf{B} \boxminus \mathbf{B} \boxminus \mathbf{B}, \\ \nabla_{\mathbf{C}} \mathcal{G}(\cdot) &= \nabla_{\mathbf{C}} \mathcal{F}(\cdot) + \alpha_C \star\star\star & \text{where } \star\star\star &= 2\mathbf{C} \quad \text{or} \quad 4\mathbf{C} \boxminus \mathbf{C} \boxminus \mathbf{C} \end{aligned}$$

$2\mathbf{A}$ (or $2\mathbf{B}$ or $2\mathbf{C}$) is added in the case of a l_1 -norm regularization, while $4\mathbf{A} \boxminus \mathbf{A} \boxminus \mathbf{A}$ (or $4\mathbf{B} \boxminus \mathbf{B} \boxminus \mathbf{B}$ or $4\mathbf{C} \boxminus \mathbf{C} \boxminus \mathbf{C}$) is added in the case of a l_2 -norm regularization. When the model is exact (F is assumed to be equal to the true number of fluorophores), these terms can be discarded ($\alpha_{\bullet} = 0$) as illustrated in the following subsection 4.4. In the case of overfactoring (F is assumed to be higher than the true number of fluorophores), these additional terms happen to be absolutely necessary to recover the true solution, yet, they also induce a bound on the obtained performance. This is the reason why we do not assume fixed values for α_A , α_B and α_C while iterations are run, but rather opt for a “hybrid” strategy: during the first half of the iterations α_A , α_B and α_C are fixed and in the second half they are discarded (*i.e.* set to zero).

C. The stepsize issue

The stepsize $\mu^{(k)}$ can be found in different ways: a global search with Enhanced Line Search (ELS) or an approximation by a line search method like backtracking [2][56] (which is a locally optimal step size method). ELS consists of looking for the value of μ that minimizes the degree twelve univariate polynomial $\mathcal{F}(\mathbf{P}_1 \boxtimes \mathbf{P}_2, \mathbf{P}_2 \boxtimes \mathbf{P}_3, \mathbf{P}_3 \boxtimes \mathbf{P}_1)$ where $\mathbf{P}_1 = \mathbf{A}^{(k)} + \mu \mathbf{D}_A^{(k)}$, $\mathbf{P}_2 = \mathbf{B}^{(k)} + \mu \mathbf{D}_B^{(k)}$ and $\mathbf{P}_3 = \mathbf{C}^{(k)} + \mu \mathbf{D}_C^{(k)}$ with:

$$\mathbf{D}^{(k)} = \begin{pmatrix} \mathbf{D}_A^{(k)} \\ \mathbf{D}_B^{(k)} \\ \mathbf{D}_C^{(k)} \end{pmatrix}, \quad \mathbf{d}^{(k)} = \begin{pmatrix} \text{vec}\{\mathbf{D}_A^{(k)}\} \\ \text{vec}\{\mathbf{D}_B^{(k)}\} \\ \text{vec}\{\mathbf{D}_C^{(k)}\} \end{pmatrix} = \begin{pmatrix} \mathbf{d}_A^{(k)} \\ \mathbf{d}_B^{(k)} \\ \mathbf{d}_C^{(k)} \end{pmatrix} \quad (14)$$

We can choose to use ELS at each iteration or backtracking alternating with ELS (in this case, the algorithm is less time consuming since ELS is only performed every 10 or 20 iterations for example). In the simulations that are presented in the following, we will only report the results that are obtained using ELS at each iteration. The algorithm will be denoted by NTF-HCGP_{ELS} (for Nonnegative Tensor Factorization using a Hadamard parametrization, a non linear Conjugate Gradient optimization algorithm, a Penalized modified objective function and an Enhanced Line Search).

D. Numerical simulations on synthetic data

The purpose of this subsection is to evaluate the performance of the NTF-HCGP_{ELS} algorithm only, regardless of the problems that are tied to the preprocessing. But we also want to stick to the experimental example presented in the next section that is why a synthetic data tensor is built as follows: emission and excitation spectra of three fluorophores, in our case tyrosine, phenylalanine and tryptophan have been downloaded at the following address: <http://omlc.org/spectra/PhotochemCAD/index.html>. The three resulting Fluorescence Excitation Emission Matrices, $\mathbf{a}_i \mathbf{b}_i^T$, $\forall i = 1, \dots, 3$ are 63×58 matrices. A 8×3 random nonnegative matrix \mathbf{C} has been generated too and a nonnegative $63 \times 58 \times 8$ tensor \mathbf{T} has been simulated.

On Fig. 2, we have given the reference emission and excitation spectra as well as those that are estimated using the NTF-HCGP_{ELS} starting from a random initialisation and without regularization terms ($\alpha_A = \alpha_B = \alpha_C = 0$)).

We can observe that in the case of a perfect noiseless trilinear model and a tensor of known rank, both spectra and FEEM are perfectly recovered with the NTF-HCGP_{ELS} algorithm. On this example, similar results would be obtained thanks to other classical CP decomposition algorithms such as the NTF-HALS or NTF-ALS algorithms [14], or Bro's Parafac algorithm [3][5]. The same fit between results would be observed on the FEEM.

Yet, to be able to establish a real comparison between the different algorithms, we now introduce two error

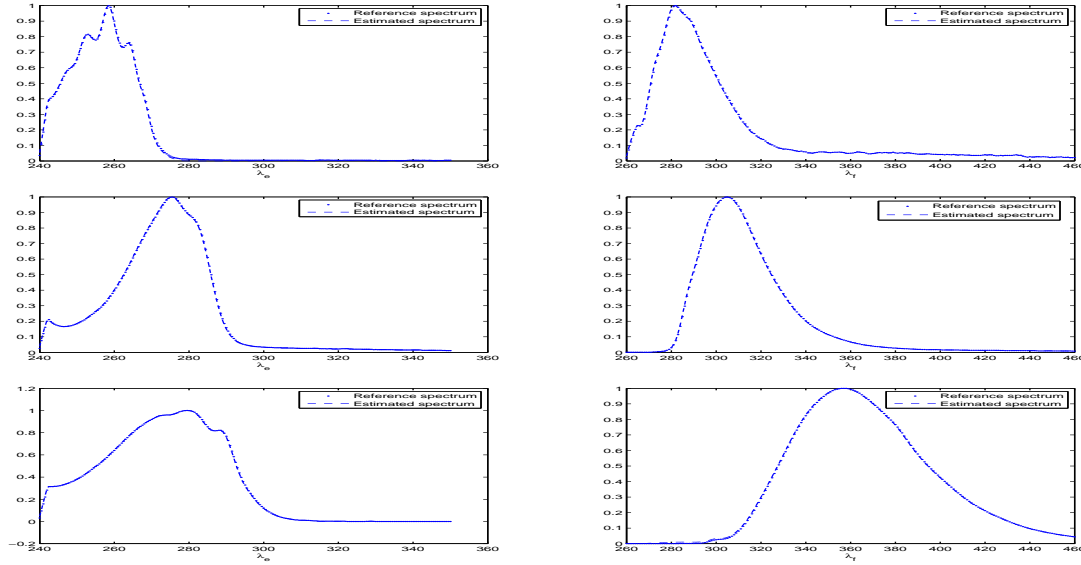


Fig. 2. Normalized estimated spectra. Left: excitation spectra, right emission spectra. First line: phenylalanine, second line: tyrosine and third line: tryptophan. Spectrum estimated by nonnegative CP decomposition (dashed line), reference spectrum collected in [http://omlc.org.edu/spectra/Photochem CAD/index.html](http://omlc.org.edu/spectra/PhotochemCAD/index.html) (dotted line). In this case, the rank is assumed known and equal to 3

indices. We have chosen to use the following ones:

$$E_1 = \|\mathbf{T} - \hat{\mathbf{T}}\|_F^2, \quad E_{1\text{dB}} = 10 \log_{10}(E_1), \quad (15)$$

with $\hat{\mathbf{T}} = \sum_{f=1}^F \hat{\mathbf{a}}_f \otimes \hat{\mathbf{b}}_f \otimes \hat{\mathbf{c}}_f$ and $\hat{\mathbf{a}}$, $\hat{\mathbf{b}}$ and $\hat{\mathbf{c}}$ the estimated factors. This error index measures the reconstruction error and should tend to 0 ($-\infty$ in logarithmic scale). The second performance index (which cannot be used on experimental data) measures the distance between the FEEM of actual and estimated components, and the best results are obtained when E_2 is found to be close to 0 in linear scale ($-\infty$ in logarithmic scale):

$$E_{2\text{dB}} = 10 \log_{10} \left(\sum_{i=1}^F \|\mathbf{a}_i \mathbf{b}_i^T - \hat{\mathbf{a}}_i \hat{\mathbf{b}}_i^T\|_1 \right) \quad (16)$$

But first, to be able to apply this second error index and thus to compare the images corresponding to the same organic component, actual and estimated FEEM are normalized and, then, they are sorted since permutations between the columns of the factors can occur (we have to find among the $F!$ possible permutations of the F columns of the estimated loading matrices, the one that leads to the smaller distances between the estimated FEEM and the actual FEEM). On the Fig. 3, the results obtained with different algorithms (for instance: the NTF – HCGP_{ELS}, the NTF – ALS, NTF – HALS and Bro's N – way algorithms) are compared in the case of an exact model. The evolution of the performance indices are given versus the iterations.

We clearly observe on Fig. 3 that in the case of a perfect trilinear model and a known rank, the four algorithms are able to recover the true solution. Moreover, the NTF – HCGP_{ELS} is the most computer time consuming, which is not surprising according to algorithmic costs provided in [56]. We notice that performances are bounded for the NTF – ALS algorithm: this is due to the use of fixed regularization terms.

Secondly, the robustness versus model errors is evaluated: the rank is now overestimated. We still consider the same tensor \mathbf{T} , but, since F is generally unknown, we try to perform the CP decomposition with the assumption that $\hat{F} = 4$. Regarding the NTF – HCGP_{ELS}, regularization terms are now used. We choose the same $\alpha_A = \alpha_B = \alpha_C = 10^{-3}$ and l_1 -norm regularization terms for each factors during the first half of the iterations, then in the second half they are discarded to avoid a bound in the performances. We choose l_1 -norm regularization terms since it was shown in [67] that they lead to the best results. The same random initializations are used for all the algorithms. The five resulting FEEM as well as the FEEM of reference are displayed on Fig. 4. On its top-left, we have given the FEEM of reference, on its top-right those that are estimated using the NTF-HALS algorithm, on its middle-left using the NTF-ALS algorithm, on its middle-right using Bro’s Nway algorithm, below-left, using the NTF-HCGP_{ELS} without regularization terms ($\alpha_A = \alpha_B = \alpha_C = 0$), finally, below-right, those that are estimated thanks to the NTF-HCGP_{ELS} using regularization terms ($\alpha_A = \alpha_B = \alpha_C = 10^{-3}$ and l_1 -norm). Ideally, unnecessary components should be equal to 0. As it can be observed on the Figure 4, it is more likely the case using the improved method that we have introduced: only small residues are visible on this set of images, but to reach that result, we also notice that it was necessary to introduce regularization terms. Without regularization terms, the algorithm was not able to recover the exact solution (one of the compound that should be estimated is broken into two parts).

In the Fig. 5, we compare the results obtained thanks to the four algorithms, using the two aforementioned error indices. We now clearly observe on Fig. 5 that, in this case, the NTF – HCGP_{ELS} outperforms the three other algorithms but only when regularization terms are used. We can also observe on this same chart that when regularization terms are used, performances reach a plateau, then, by canceling those regularization terms the error indice can decrease again.

Finally, on Fig. 6, we also illustrate the impact of the chosen initialisation on the different algorithms. To that aim, 60 different random initializations were generated according to a uniform law. The results obtained thanks to the different CP decomposition algorithms are compared on this chart: performances are sorted in descending order from the best performance to the worst. We observe that in 72% of the cases, our algorithm was able to outperform other methods (-45dB versus -10dB). In the 28% other cases, we reach the same level of performance as the one obtained with non regularized algorithms (which means that when the random initial solution is chosen really far from the solution, either more iterations should be used or the values for the regularization terms should be modified).

	Main peak	
	Excitation	Emission
Phenylalanine	257	279-282
Tyrosine	275	310
Tryptophan	275	340-350

TABLE I

FLUORESCENCE MAIN PEAK POSITION ACCORDING TO [HTTP://OMLC.ORG.L/EDU/SPECTRA/PHOTOCHEM/CAD/INDEX.HTML](http://omlc.orgl.edu/spectra/PhotochemCAD/index.html).

V. COMPUTER SIMULATIONS ON EXPERIMENTAL FLUORESCENCE SPECTROSCOPY DATA

To better assess pros and cons of the suggested approach within the fluorescence spectroscopy framework and to be able to compare it to more conventional approaches of the literature, a challenging case has been prepared in the PROTEE laboratory: it consists of a very small number of FEEM since we use three samples containing each a mixing of three known fluorophores. Moreover, some components have close spectra and scatter peaks strongly overlap the fluorescence signal. In this example like in section IV.D, the considered fluorophores are still tyrosine, phenylalanine and tryptophan. Their respective proportions in each sample are also known and they are equal to $[2/1/1]$, $[1/1/2]$ and $[1/2/1]$. The expected positions of the fluorophores' main peak are also recalled in Table I (these information have been collected from the database which can be accessed via the following web address <http://omlc.orgl.edu/spectra/PhotochemCAD/index.html>). The purpose of this experimental example is not to directly apply blindly the suggested approach on large environmental data sets and show the obtained results but to illustrate the fact that it remains really important to correctly calibrate the whole processing chain on known experimental mixtures in order to be able to choose all the parameters involved in a given approach. Whatever the considered method, a bad choice of parameters can really alter the obtained results. Besides the impact of the preprocessing on the obtained results should not be underestimated. It is only after this calibration stage that it becomes possible to use such an approach in an unsupervised or blind mode. The purpose of this example is also to show that this approach can still operate in difficult cases where the separability limit is nearly reached.

In this example, the considered 3-way tensor is of size $101 \times 47 \times 3$ *i.e.* $\mathbf{T} \in \mathbb{R}^{101,47,y}$ with $y = 3$. The low-rank constraint implies that the rank, say r , of the matrix $\mathbf{T}_{(3)}$ obtained by unfolding the tensor in the mode-3 is constrained by y , *i.e.* in this example $r \leq 3$.

Five different scenarios are then considered:

- in the first one, Raman and Rayleigh scattering peaks are eliminated thanks to the suggested preprocessing

based on the coupling of the mathematical morphology filtering technique and the Zepp's method (Eq. (8) with $\alpha = 0.3$). The nonnegative CP decomposition algorithm described in Section V (NTF – HCGP_{ELS}) is then applied.

- in the second scenario, Raman and Rayleigh scattering peaks are eliminated thanks to the Zepp's method only. The same NNCP decomposition algorithm as in the first scenario is used with exactly the same initialisation and regularization terms.
- in the third scenario, we have used the same preprocessing as in the first scenario, but the CP decomposition algorithm that is used is now the one suggested by R. Bro in [3], based on an Alternating Least Square (ALS) algorithm.
- in the fourth scenario, we have used a detector based on the amplitude to detect scatters, then data corresponding to high values are considered as missing data *i.e.* the tensor of weights \mathbf{W} can be built and the approach suggested in [65] can be used (this “marginalization” approach consists of ignoring missing data).
- in the fifth scenario, we have used the first stages of the Zepp method: the position and spread of the scattering peaks are modeled thanks to a set of 4 second degree equations (the same parameters as those used in scenario 1 and 2 are used). Then, instead of excising the scattering peaks in the scan data, the found positions are exploited to build the tensor of weights \mathbf{W} , and the FEEM are directly processed using missing data *i.e.* thanks to the same algorithm as in the fourth scenario.

All the estimated (normalized) emission resp. excitation spectra are displayed in Fig. (7). We clearly observe the fact that the preprocessing can significantly modify/alter the obtained results: a noticeable shift in the estimated main peak position or a net widening of the spectrum can occur. This phenomenon is emphasized in the scenario 2, for the tyrosine (whether the emission or excitation spectrum is considered), but also for the phenylalanine (on its excitation spectrum), when only the Zepp's method is used to preprocess the FEEM. In this example the spread value of one of the four scatters is just chosen a little too high and since the Raman and Rayleigh scatters are overlapping with the fluorescence peaks for these two fluorophores, so, when the scatters were eliminated, a part of the fluorescence peaks were eliminated too inducing either a shift or a spreading of the spectrum. We also observe that results are rather deteriorated than improved if scatters are handled by setting them to missing values (scenarios 4 and 5). When the same coupled preprocessing is applied to the FEEM (scenarios 1 and 3), we can observe that the estimated peak positions are nearly the same whether the nonnegative CP decomposition that we suggest or the classical Bro's CP decomposition algorithm is used. However, the shape of the estimated spectra are not the same with both methods. Spectra exhibiting negative values can be obtained with the Bro's algorithm. In this example, the best results are observed in the first scenario. Yet, if emission spectra are quite close to the reference spectra, it is nevertheless less true for the excitation spectra. We notice

a small peak shift (around a 3 nm gap for phenylalanine and 3 nm for tryptophan), and a broadening of the excitation spectrum (and to a lesser extend for the emission spectrum) for tyrosine and phenylalanine. The tryptophan main fluorescence peak is found in (272 nm; 354 nm), (275 nm; 305 nm) for tyrosine and (254 nm; 281 nm) for phenylalanine which fits quite well with the information collected in Tab. I. This example clearly emphasizes the impact of the preprocessing on the obtained results and the interest one can have in using an efficient nonnegative CP decomposition algorithm instead of an unconstrained CP decomposition algorithm.

VI. CONCLUSION

In this article, focusing on the fluorescence spectroscopy framework, we have presented a novel approach to handle the preprocessing of FEEM and the nonnegative CP decomposition of the resulting three-way tensor of FEEM. Computer simulations have been provided to enlighten the effectiveness and the robustness of the whole processing chain (even in really difficult cases where the Raman and Rayleigh scatters overlap with fluorophores' fluorescence peaks). The method has been compared with other existing approaches. Furthermore, this study shows that the calibration of the whole processing chain on known but experimental mixtures remains an important stage since it enables to fix the different parameters involved in a given method. Whatever the considered approach - improper settings can really alter the obtained results. Hence, the impact of the preprocessing on the obtained results should not be underestimated. Concerning the NNCP decomposition algorithm, the critical importance of additional regularization terms (that will be progressively decreased or canceled after the first half of iterations) has also been emphasized when the number of fluorophores is overestimated.

REFERENCES

- [1] C. Belzile and L. Guo, "Optical properties of low molecular weight and colloidal organic matter: Application of the ultrafiltration permeation model to dom absorption and fluorescence," *Marine Chemistry*, vol. 98, no. 2-4, pp. 183–196, 2006.
- [2] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, March 2004.
- [3] R. Bro, "Parafac: tutorial and applications," *Chemometr. Intell. Lab.*, vol. 38, pp. 149–171, 1997.
- [4] —, "Multi-way analysis in the food industry: models, algorithms and applications," Ph.D. dissertation, University of Amsterdam, Amsterdam, The Netherlands, 1998.
- [5] R. Bro and C. A. Andersson, *N-Way Toolbox for Matlab*, <http://www.models.life.ku.dk/nwaytoolbox>.
- [6] R. Bro and S. D. Jong, "A fast non-negativity-constrained least squares algorithm," *Journal of Chemometrics*, vol. 11, no. 5, pp. 393–401, 1997.
- [7] R. Bro and H. A. Kiers, "A new efficient method for determining the number of components in parafac models," *Journal of chemometrics*, vol. 17, pp. 274–286, 2003.
- [8] J. B. C. Bugden, C. W. Yeung, P. E. Kepkay, and K. Lee, "Application of ultraviolet fluorometry and excitation-emission matrix spectroscopy (eems) to fingerprint oil and chemically dispersed oil in seawater," *Marine pollution bulletin*, vol. 56, no. 4, pp. 677–685, 2008.

- [9] D. Bunker, L. Han, and S. Zhang, "A proximal anls algorithm for nonnegative tensor factorization with a periodic enhanced line search," *Application of Mathematics*, vol. 58, no. 5, pp. 493–509, October 2013.
- [10] P. Carroll and J. J. Chang, "Analysis of individual differences in multi-dimensional scaling via n-way generalization of eckart-young decomposition," *Psychometrika*, vol. 35, pp. 283–319, 1970.
- [11] E. M. Carstea, A. Baker, M. Bieroza, and D. Reynolds, "Continuous fluorescence excitation-emission matrix monitoring of river organic matter," *Water Research*, vol. 44, no. 18, pp. 5356–5366, 2010, doi:10.1016/j.watres.2010.06.036.
- [12] E. Ceulemans and H. A. Kiers, "Selecting among three-mode principal component models of different types and complexity: a numerical convex hull based method," *British journal of mathematical and statistical psychology*, vol. 59, no. 1, pp. 133–150, 2006.
- [13] Z.-P. Chen, H.-L. Wu, J.-H. Jiang, Y. Li, and R.-Q. Yu, "A novel trilinear decomposition algorithm for second-order linear calibration," *Chemometrics and Intelligent Laboratory Systems*, vol. 1, no. 52, pp. 75–86, January 2000.
- [14] A. Cichocki, R. Zdunek, A. H. Phan, and S. I. Amari, *Non negative matrix and tensor factorizations: Application to exploratory multi-way data analysis and blind separation*. Wiley, 2009.
- [15] P. G. Coble, "Characterization of marine and terrestrial dom in seawater using excitation-emission matrix spectroscopy," *Marine Chemistry*, vol. 52, pp. 325–346, 1996.
- [16] J. Coloigner, A. Karfoul, L. Albera, and P. Comon, "Line search and trust region strategies for canonical decomposition of semi-nonnegative semi-symmetric 3rd order tensors," *Linear algebra and its application* 450, pp. 334–374, 2014.
- [17] P. Comon, X. Luciani, and A. L. F. De Almeida, "Tensor decompositions, alternating least squares and other tales," *Journal of Chemometrics*, vol. 23, pp. 393–405, August 2009.
- [18] J. D. Costa, M. Haardt, and F. Roemer, "Robust methods based on the hosvd for estimating the model order in parafac models," in *Sensor array and multichannel signal processing workshop (SAM)*, IEEE, Ed., 2008, pp. 510–514.
- [19] N. Ferretto, M. Tedetti, C. Guigue, S. Mounier, R. Redon, and M. Goutx, "Identification and quantification of known polycyclic aromatic hydrocarbons and pesticides in complex mixtures using fluorescence excitationemission matrices and parallel factor analysis," *Source Chemosphere*, no. 107, pp. 344–353, 2014.
- [20] A. Franc, "Etude algébrique des multi-tableaux : apport de l'algèbre tensorielle," PhD thesis, University of Montpellier II, Montpellier, France, 1992.
- [21] A. G. Frenich, D. P. Zamora, J. M. Vidal, and M. M. Galera, "Resolution and quantitation of mixtures with overlapped spectra by orthogonal projection approach and alternating least squares," *Analytica Chimica Acta*, vol. 449, pp. 143–155, 2001.
- [22] C. Gobinet, E. Perrin, and R. Huez, "Application of nonnegative matrix factorization to fluorescence spectroscopy," in *EUSIPCO*, Vienna, Austria, September 2004.
- [23] R. Gonzalez and R. Woods, *Digital image processing*, 2nd ed., P. E. International, Ed. Prentice Hall, 2001.
- [24] R. A. Harshman, "Foundation of the Parafac procedure: models and conditions for an explanatory multimodal factor analysis," *UCLA Working papers in phonetics*, vol. 16, pp. 1–84, 1970.
- [25] R. A. Harshman and W. S. D. Sarbo, "An application of parafac to a small sample problem, demonstrating preprocessing, orthogonality constraints, and split-half diagnostic techniques," *Statistique et Analyse des Données*, vol. 13, pp. 14–32, 1984.
- [26] Z. He, A. Cichocki, S. Xie, and K. Choi, "Detecting the number of clusters in n-way probabilistic clustering," *IEEE transaaction on pattern analysis and machine intelligence*, vol. 32, no. 11, pp. 2006–2021, 2010.
- [27] F. L. Hitchcock, "The expression of a tensor or a polyadic as a sum of products," *J. Math. and Phys.*, vol. 6, pp. 165–189, 1927.
- [28] L.-Q. Hu, H.-L. Wu, J.-H. Jiang, Q.-J. Han, A.-L. Xia, and R.-Q. Yu, "Estimating the chemical rank of three-way data arrays by a simple linear transform incorporating monte carlo simulation," *Talanta*, vol. 71, pp. 373–380, 2007.
- [29] L. Ingber, "Simulated annealing: Practice versus theory," *Mathematical Computer Modelling*, vol. 18, no. 11, pp. 29–57, 1993.
- [30] J. H. Jiang, H. L. Wu, Y. Li, and R. Q. Yu, "Alternating coupled vectors resolution (acover) method for trilinear analysis of three-way data," *Journal of Chemometrics*, vol. 13, no. 6, pp. 557–578, 1999.
- [31] R. D. Jiji, G. G. Andersson, and K. S. Booksh, "Applicaton of parafac for calibration with excitation-emission matrix fluorescence spectra of three classes of environmental pollutants," *J. of Chemometrics*, vol. 14, no. 3, pp. 171–185, 2000.

- [32] C. G. Khatri and C. R. Rao, "Solutions to some functional equations and their applications to characterization of probability distributions," *Sankhya: Indian J. Statistics*, vol. Series A, 30, pp. 167–180, 1968.
- [33] H. A. L. Kiers, "Towards a standardized notation and terminology in multiway analysis," *J. Chemometrics*, pp. 105–122, 2000.
- [34] J. Kim and H. Park, "Fast nonnegative tensor factorization with an active-set-like method," *High-Performance Scientific Computing: Algorithms and Applications*, Springer, pp. 311–326, 2012.
- [35] F. Kirly and A. Ziehe, "Approximate rank-detecting factorization of low-rank tensors," in *International Conference on Acoustic Speech and Signal Processing (ICASSP'2013)*, Vancouver, Canada, 2013.
- [36] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *Siam Review*, vol. 51, no. 3, pp. 455–500, September 2009.
- [37] J. R. Lacowicz, *Principle of fluorescence spectroscopy*, K. Academic, Ed. Plenum Publishers, 1999.
- [38] L. H. Lim and P. Comon, "Nonnegative approximations of nonnegative tensors," *Journal of Chemometrics*, vol. 23, pp. 432–441, August 2009.
- [39] X. Luciani and L. Albera, "Canonical polyadic decomposition based on joint eigenvalue decomposition," *Chemometrics and Intelligent Laboratory Systems*, vol. 132, pp. 152–167, March 2014.
- [40] X. Luciani, S. Mounier, R. Redon, and A. Bois, "A simple correction method of inner filter effects affecting FEEM and its application to the Parafac decomposition," *Chemometrics and Intel. Lab. Syst.*, vol. 96, no. 2, pp. 227–238, 2009.
- [41] D. G. Luenberger and Y. Ye, *Linear and non linear programming*, 3rd ed. Wiley, 2008.
- [42] E. Malinowski, "Window factor analysis: theoretical derivation and application to flow injection analysis data," *Journal of Chemometrics*, vol. 6, pp. 29–40, 1992.
- [43] S. Mehrotra, "On the implementation of a primal-dual interior point method," *SIAM Journal on Optimization*, vol. 2, pp. 575–601, 1992.
- [44] J. Nelder and R. Mead, "A simplex method for function minimization," *The Computer Journal*, vol. 7, pp. 308–313, 1965.
- [45] J. Nocedal and S. Wright, *Numerical Optimization*, Springer, Ed., New York, NY, 1999.
- [46] P. Paatero, "A weighted non-negative least squares algorithm for three-way Parafac factor analysis," *Chemometrics Intell. Lab. Systems*, vol. 38, no. 2, pp. 223–242, 1997.
- [47] —, "Construction and analysis of degenerate Parafac models," *Journal of Chemometrics*, vol. 14, no. 3, pp. 285–299, 2000.
- [48] M. L. Pascu, D. Cristu, A. Vasile, and N. Moise, "Pesticide monitoring in water using laser-induced fluorescence and absorption techniques," in *Proc of ROMOPTO'94, SPIE Vol. 2461*, Bucharest, Romania, 1995, pp. 648–652.
- [49] D. Patra and A. Mishra, "Recent developments in multi-component synchronous fluorescence scan analysis," *TrAC Trends in Analytical Chemistry*, vol. 21, no. 12, pp. 787–798, 2002.
- [50] T. Persson and M. Wedborg, "Multivariate evaluation of the fluorescence of aquatic organic matter," *Anal. Chim. Acta* 434, p. 179 192, 2001.
- [51] E. Polak, *Optimization algorithms and consistent approximations*. Springer, 1997.
- [52] K. Price and R. Storn, "Differential evolution," *Dr. Dobbs's Journal*, vol. 264, pp. 18–24, 1997.
- [53] D. M. Reynolds, "The differentiation of biodegradable and non-biodegradable dissolved organic matter in wastewaters using fluorescence spectroscopy," *J. of Chemical Technology and Biotechnology*, vol. 8, no. 965-972, 77.
- [54] F. Roemer and M. Haardt, "A closed-form solution for parallel factor (parafac) analysis," in *IEEE international conference on acoustics, speech, and signal processing*, 2008, p. 23652368.
- [55] J.-P. Royer, P. Comon, and N. Thirion-Moreau, "Computing the nonnegative 3-way tensor factorization using tikhonov regularization," in *International Conference on Acoustic Speech and Signal Processing (ICASSP'2011)*, Prague, Czech Republic, May 2011, pp. 2732–2735.
- [56] J.-P. Royer, N. Thirion-Moreau, and P. Comon, "Computing the polyadic decomposition of nonnegative third order tensors," *Eurasip Signal Processing*, vol. 91, no. 9, pp. 2159–2171, September 2011.
- [57] —, "Nonnegative 3-way tensor factorization taking into account possible missing data," in *European Signal Processing Conference (EUSIPCO'2012)*, Bucharest, Romania, August 2012.

- [58] J. Serra, *Image Analysis and Mathematical Morphology*. New York, USA: Academic Press, 1982.
- [59] —, *Image Analysis and Mathematical Morphology*, vol. 2. New York, USA: Academic Press, 1988.
- [60] A. Smilde, R. Bro, and P. Geladi, *Multi-Way Analysis with applications in the chemical sciences*. Wiley, 2004.
- [61] C. A. Stedmon, S. Markager, and R. Bro, “Tracing dissolved organic matter in aquatic environments using a new approach to fluorescence spectroscopy,” *Marine Chemistry*, vol. 82, pp. 239–254, 2003.
- [62] S. R. Sternberg, “Grayscale morphology,” *Computer Vision, Graphics, and Image Processing*, vol. 35, pp. 333–355, 1986.
- [63] P. Tichavsky and Z. Koldovsky, “Simultaneous search for all modes in multilinear models,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP’10)*, 2010.
- [64] M. E. Timmerman and H. A. Kiers, “Three-mode principal components analysis: choosing the number of components and sensitivity to local optima,” *British journal of mathematical and statistical psychology*, vol. 53, no. 1, pp. 1–16, 2000.
- [65] G. Tomasi and R. Bro, “Parafac and missing values,” *Chemometrics Intell. Lab. Systems*, 2004.
- [66] H. Vandenhove, C. Hurtgen, and T. Payne, “Radionuclides in the environment: uranium,” *Radionuclides in the environment*, pp. 261–272, May 2010, west Sussex, United Kingdom, John Wiley and Sons.
- [67] X. T. Vu, C. Chau, S. Maire, and N. Thirion-Moreau, “Study of different strategies for the canonical polyadic decomposition of nonnegative third order tensors with application to the separation of spectra in 3d fluorescence spectroscopy,” in *IEEE Machine Learning for Signal Processing (MLSP’14)*, Reims, France, September 2014.
- [68] Y.-J. Yu, H.-L. Wu, J.-F. Nie, S.-R. Zhang, S.-F. Li, Y.-N. Li, S.-H. Zhhou, and R.-Q. Yu, “A comparison of several trilinear second-order calibration algorithms,” *Chemometrics and Intelligent Laboratory Systems*, vol. 106, no. 1, pp. 93–107, 2011.
- [69] R. G. Zepp, W. M. Sheldon, and M. A. Moran, “Dissolved organic fluorophores in southeastern us coastal waters: correction method for eliminating rayleigh and raman scattering peaks in excitation-emission matrices,” *Marine Chemistry*, vol. 89, no. 1-4, pp. 15–36, 2004.
- [70] B. Zhu, S. A. Pennell, and D. K. Ryan, “Characterizing the interaction between uranyl ion and soil fulvic acid using parallel factor analysis and a two-site fluorescence quenching model,” *Microchemical Journal*, vol. 115, pp. 51–57, 2014.

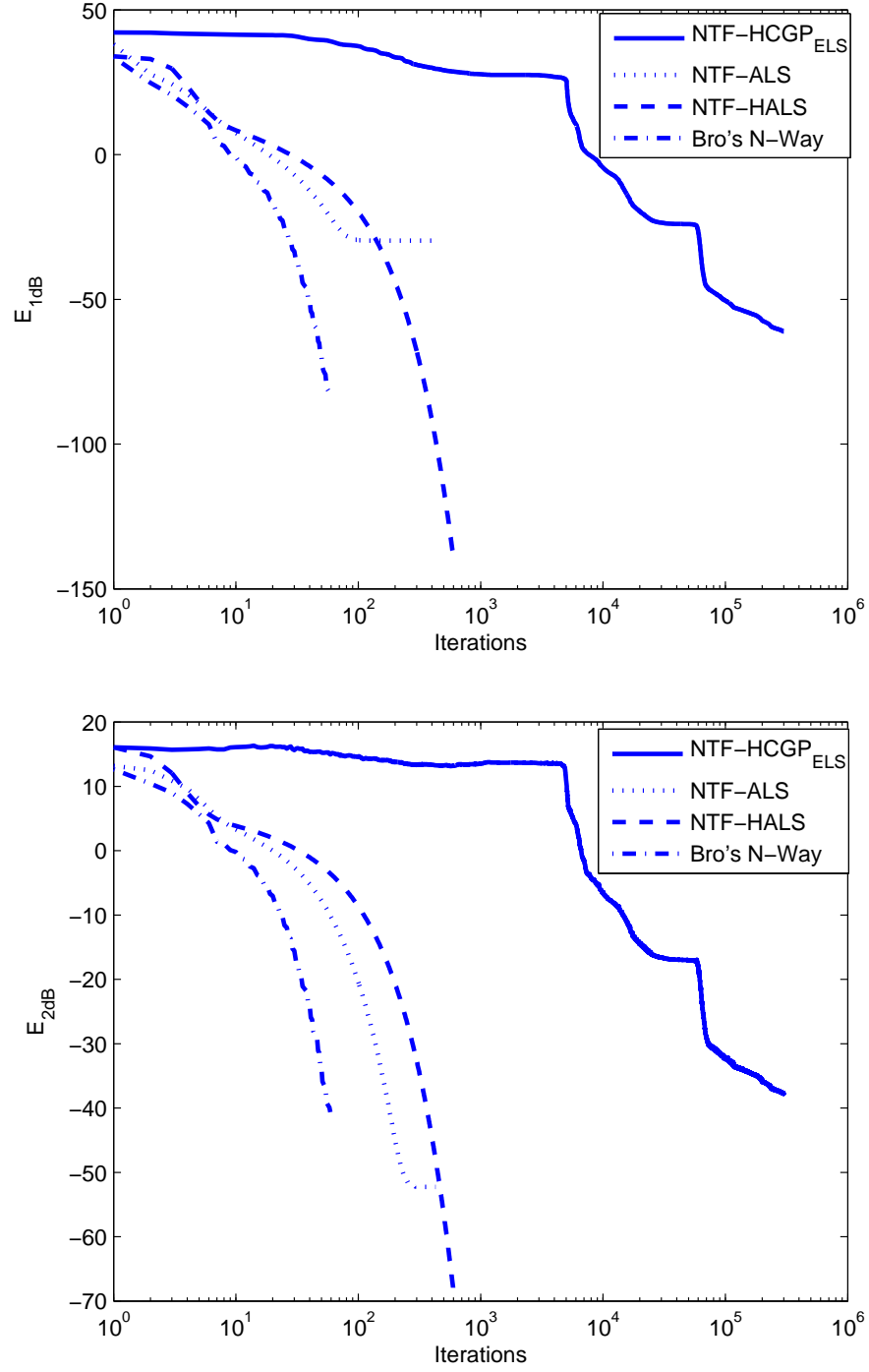


Fig. 3. Top: performance index E_{1dB} versus iteration (exact model) ; below: performance index E_{2dB} versus iteration (exact model)

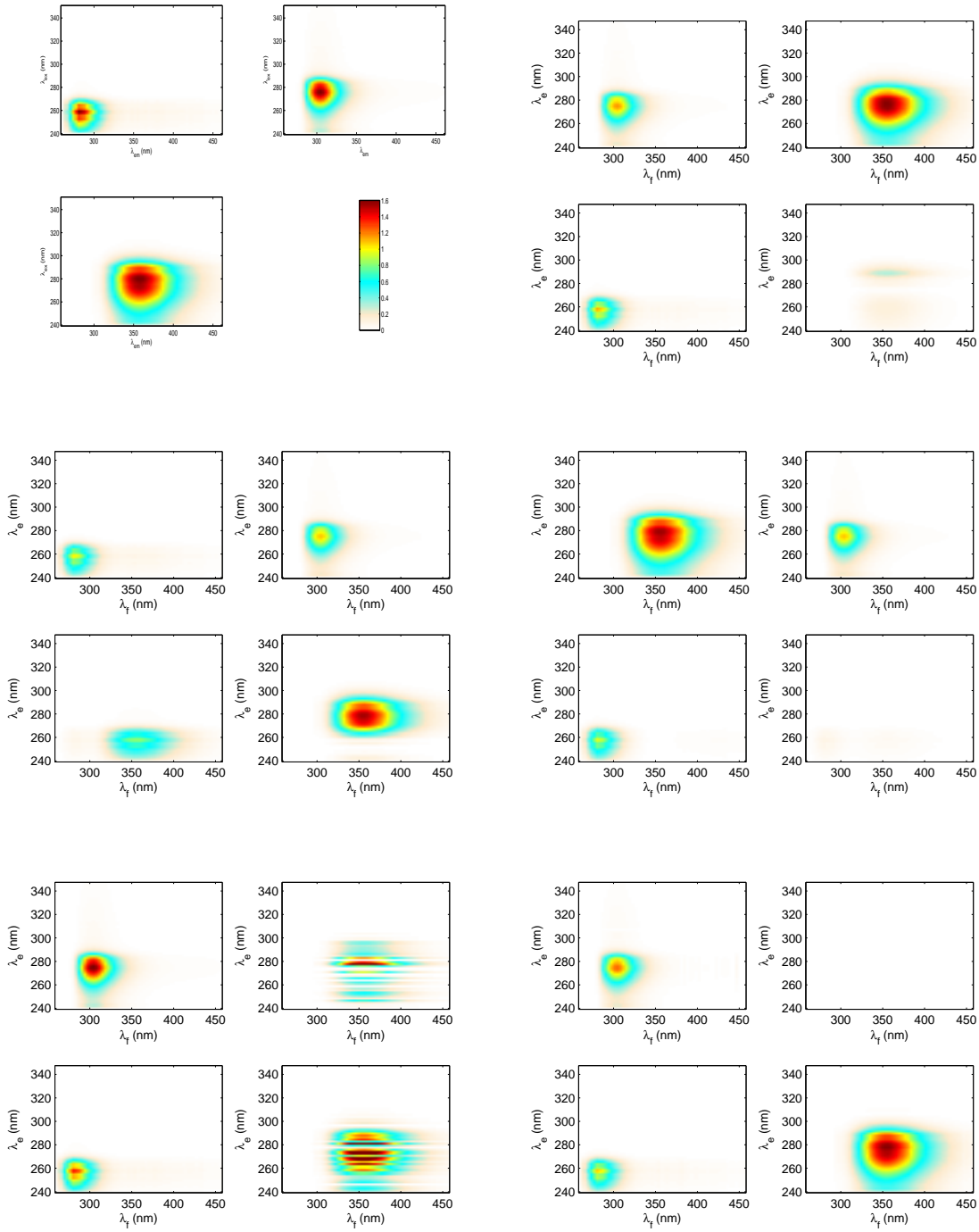


Fig. 4. The 3 FEEM of reference (top, left), the estimated FEEM: NTF-HALS (top, right), NTF-ALS (middle, left), Bro's N -way (middle, right), NTF-HCGP_{ELS} without regularization terms (bottom, left), NTF-HCGP_{ELS} with regularization terms (bottom, right)

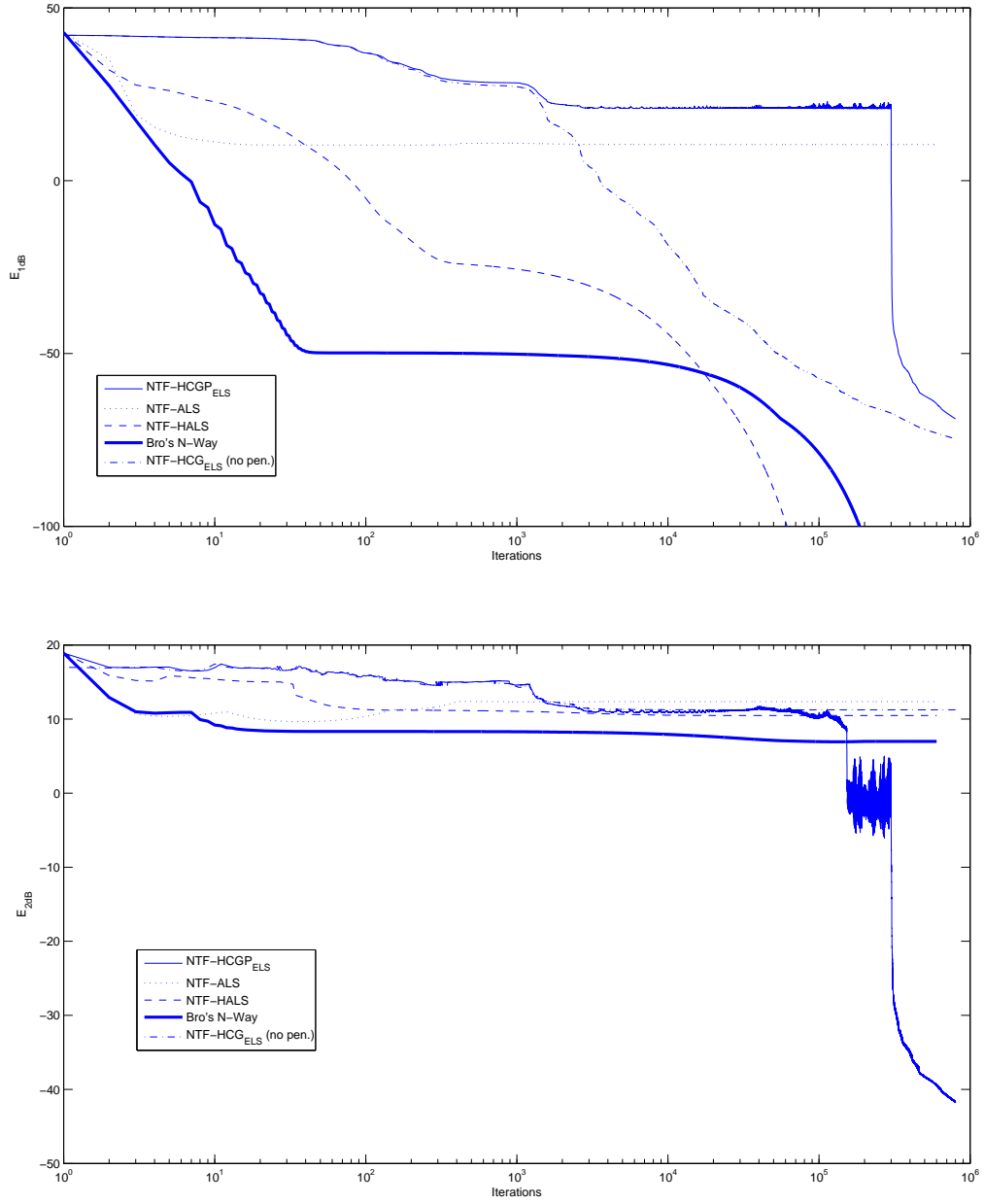


Fig. 5. Top: performance index E_{1dB} versus iteration (overestimated case) ; below: performance index E_{2dB} versus iteration (overestimated case)

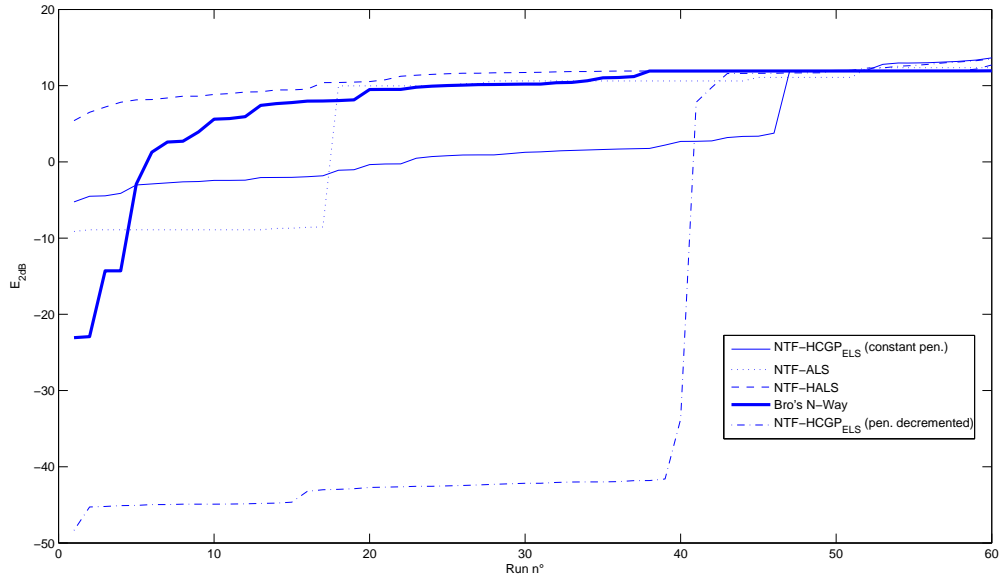


Fig. 6. Results obtained with different CP decomposition algorithms using 60 random initializations. Performances are sorted in descending order from the best performance to the worst.

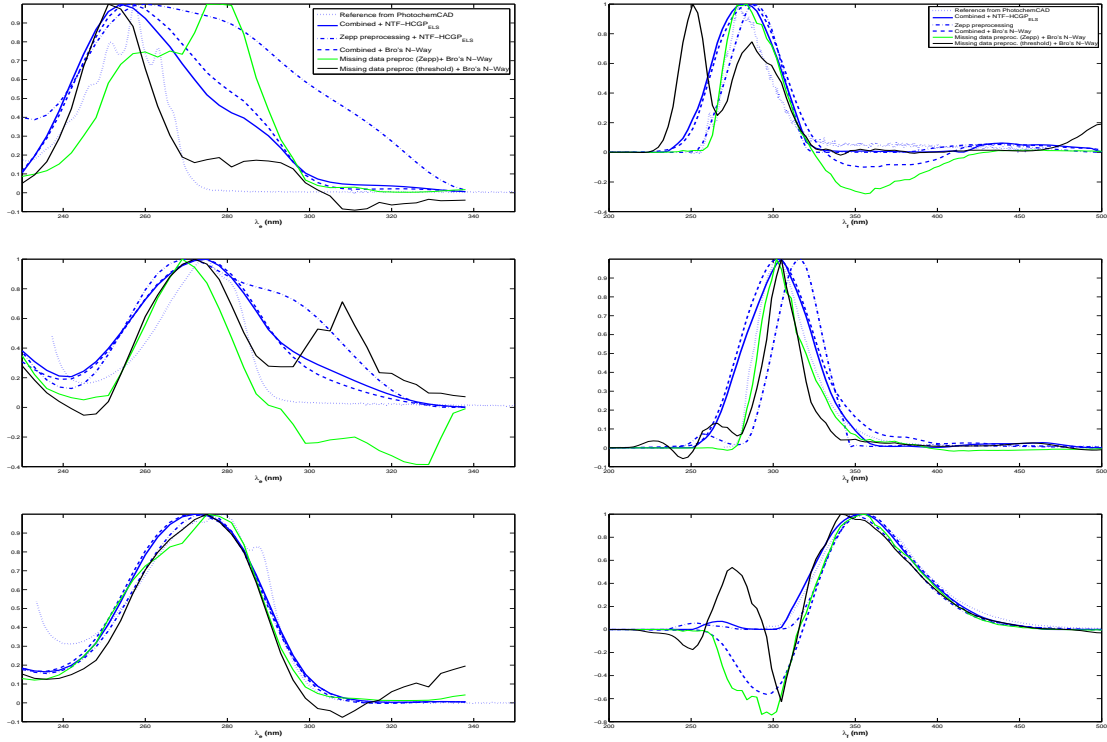


Fig. 7. Normalized estimated spectra. Left: excitation spectra, right emission spectra. First line: phenylalanine, second line: tyrosine and third line: tryptophan. Spectrum estimated by nonnegative CP decomposition + math. morphology coupled with Zepp's method preprocessing (solid line), Spectrum estimated by nonnegative CP decomposition + Zepp's method preprocessing (dashed dot line), Spectrum estimated by Bro's NWay algorithm + math. morphology coupled with Zepp's method preprocessing (dashed line), reference spectrum collected in <http://omlc.org/spectra/PhotochemCAD/index.html> (dotted line), algorithm suggested in [65] + detector of amplitude (data corresponding to high values are considered as missing data) (black line), algorithm suggested in [65] + first stages of Zepp's method to detect scatters and to handle them as missing data (green line).